A Gentle Adventure Mechanising Message Passing Concurrency Systems

An Experience/Walkthrough Report

David Castro-Perez, Lorenzo Gheri, Francisco Ferreira, Martin Vassor, and Nobuko Yoshida

DisCoTec22

Imperial College London





Roadmap

1. Binders and Linearity

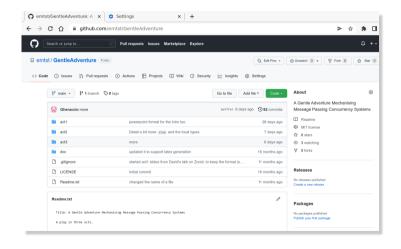
Act I: Binary Session Types

2. Multiparty Processes and Coinduction

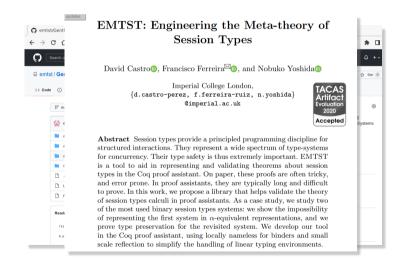
Act II: Mechanising Multiparty Processes

Act III: Mechanising Multiparty Session Types

SmolEMTST: Tutorial Repository



SmolEMTST: Tutorial Repository



SmolEMTST: Tutorial Repository

David Castro-Perez

Imperial College London and University of Kent, UK

Lorenzo Gheri

Imperial College London UK

We design and implement Zooid, a domain specific lan-

guage for certified multiparty communication, embedded

in Cog and implemented aton our mechanisation framework of asynchronous multiparty session types (the first of its

kind). Zooid provides a fully mechanised metatheory for the semantics of global and local types, and a fully verified

end-point process language that faithfully reflects the type-

level behaviours and thus inherits the global types properties

such as deadlock freedom, protocol compliance, and liveness

Abstract

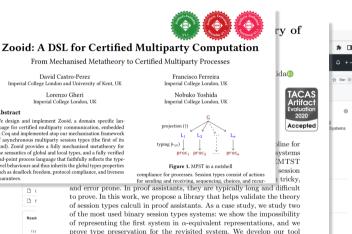
guarantees.

Bι

D) F

Reads

Tit



in the Coq proof assistant, using locally nameless for binders and small scale reflection to simplify the handling of linear typing environments.



Mechanising the Honda, Vasconcelos and Kubo's binary session type system

Honda, K., Vasconcelos, V. T., and Kubo, M. (1998). Language primitives and type discipline for structured communication-based programming. In Hankin, C., editor, <u>Programming Languages and Systems</u>, pages 122–138, Berlin, Heidelberg. Springer Berlin Heidelberg

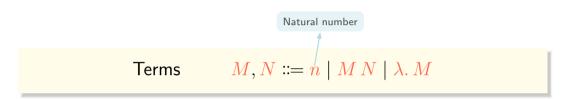
Processes: Key Features

Processes: Key Features

Binder Mechanisation: DeBruijn Indices

Terms
$$M, N := n \mid M N \mid \lambda. M$$

Binder Mechanisation: DeBruijn Indices

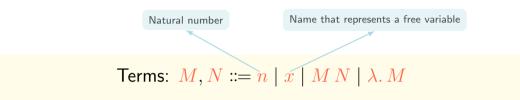


Binder Mechanisation: DeBruijn Indices

Terms $M, N := n \mid M N \mid \lambda. M$

$$\lambda$$
. $(\lambda.0)$ $(\lambda.01)$

Binder Mechanisation: Locally Nameless



Binder Mechanisation: Locally Nameless

Terms:
$$M, N := n \mid x \mid M N \mid \lambda. M$$

$$M^x \equiv \{0 o x\}M$$
 Open a term
$$\begin{split} \frac{\Gamma(x) = T}{\Gamma \vdash x : T} & \frac{\forall x \notin L \quad \Gamma, x : S \vdash M^x : T}{\Gamma \vdash \lambda. M : S \to T} \end{split}$$
 \tag{\text{\sigma} x M \in \{0 \lefta x\} M \tag{Close a term} \\ \frac{\sigma \lefta M : S \to T \quad \Gamma \tag N : S}{\Gamma \tag N : T} \\ \frac{\Gamma \tag N : S \to T \quad \Gamma \tag N : S}{\Gamma \tag N : T} \end{This identity}

Process Mechanisation

```
Inductive name : Set :=
    fnm: atom \rightarrow name
    bnm : nat \rightarrow name
Definition channel := name.
Inductive proc : Set :=
  send: channel \rightarrow exp \rightarrow proc \rightarrow proc
  receive : channel \rightarrow proc \rightarrow proc
  throw: channel \rightarrow channel \rightarrow proc \rightarrow proc
  catch : channel \rightarrow proc \rightarrow proc
 nu_ch: proc → proc (* hides a channel *)
```

```
Inductive lc:proc \rightarrow Prop:=
 lc_send : forall k e P,
     lc nm k \rightarrow
    lc_exp e \rightarrow
    1c P \rightarrow
    lc (send k e P)
| lc_receive : forall (L : seq atom) k P,
    1c \text{ nm k} \rightarrow
    (forall x, x \setminus notin L \rightarrow lc (open P x)) \rightarrow
     lc (receive k P)
lc throw: forall k k' P.
     lc_nm k \rightarrow lc_nm k' \rightarrow
    1cP→
     1c (throw k k' P)
| lc_catch : forall (L : seq atom) k P,
     1c nm k \rightarrow
     (forall x, x \setminus notin L \rightarrow lc (open P x)) \rightarrow
    ic (catch k P)
```

Semantics (Excerpt as in Paper)

R-PASS throw
$$k$$
 $[k']$; $P \mid \operatorname{catch} k$ $(k').Q \to P \mid Q$
R-COM $k ! [e]$; $P \mid k ? (x).Q \to P \mid \{x \to e\}Q$
R-CONG $P \equiv P'$ and $P' \to Q'$ and $Q' \equiv Q \Rightarrow P \to Q$
R-SCOP $P \to Q \Rightarrow \nu_c(k).P \to \nu_c(k).Q$
R-PAR $P \to P' \Rightarrow P \mid Q \to P' \mid Q$

Mechanising the Semantics

Using Locally Nameless, α -equivalent terms are syntactically equal

How do we mechanise R-PASS?

throw
$$k\left[k'\right];P\mid \mathrm{catch}\ k\left(k'\right).Q\rightarrow P\mid Q$$

Mechanising the Semantics

Using Locally Nameless, α -equivalent terms are syntactically equal

How do we mechanise R-PASS?

throw
$$k[k']; P \mid \mathsf{catch}\ k(k').Q \to P \mid Q$$

A naive attempt:

throw
$$k[k']; P \mid \text{catch } k().Q \rightarrow P \mid Q^{k'}$$

Mechanising the Semantics

Using Locally Nameless, α -equivalent terms are syntactically equal

How do we mechanise R-PASS?

throw
$$k[k']; P \mid \mathsf{catch}\ k(k').Q \to P \mid Q$$

A **WRONG** attempt:

throw
$$k[k']; P \mid \text{catch } k().Q \rightarrow P \mid Q^{k'}$$

Why is our mechanisation of r-pass wrong?

Type
$$\alpha, \beta ::= ![S]; \alpha \mid ?[S]; \alpha \mid \text{end} \mid \bot$$

Typing $\Delta ::= \cdot \mid \Delta, k : \alpha$

Subject Reduction: if $\Gamma \vdash P \triangleright \Delta$ with Δ balanced, and $P \rightarrow Q$, then there exists Δ' s.t. $\Gamma \vdash Q \triangleright \Delta'$, with Δ' balanced.

Why is our mechanisation of r-pass wrong?

Type
$$\alpha, \beta ::= ![S]; \alpha \mid ?[S]; \alpha \mid \text{end} \mid \bot$$

Typing $\Delta ::= \cdot \mid \Delta, k : \alpha$

Subject Reduction: if $\Gamma \vdash P \triangleright \Delta$ with Δ balanced, and $P \rightarrow Q$, then there exists Δ' s.t. $\Gamma \vdash Q \triangleright \Delta'$, with Δ' balanced.

Why is our mechanisation of r-pass wrong?

$$\begin{array}{ll} \mathsf{Type} & \alpha,\beta & ::= & ![S];\alpha \mid ?[S];\alpha \mid \mathsf{end} \mid \bot \\ \mathsf{Typing} & \Delta & ::= & \cdot \mid \Delta,k:\alpha \end{array}$$

Subject Reduction: if $\Gamma \vdash P \triangleright \Delta$ with Δ balanced, and $P \rightarrow Q$, then there exists Δ' s.t. $\Gamma \vdash Q \triangleright \Delta'$, with Δ' balanced.

Rule throw $k[k']; P \mid \text{catch } k().Q \rightarrow P \mid Q^{k'}$ breaks subject reduction.

The Problem with Equating α -equivalent Terms

The idea behind R-PASS:

throw
$$k[k_0]; P \mid \text{catch } k(k_1).Q \rightarrow P \mid Q'$$

if $\text{catch } k(k_1).Q \equiv_{\alpha} \text{catch } k(k_0).Q'$

The Problem with Equating α -equivalent Terms

The idea behind R-PASS:

throw
$$k[k_0]; P \mid \text{catch } k(k_1).Q \rightarrow P \mid Q'$$

if $\text{catch } k(k_1).Q \equiv_{\alpha} \text{catch } k(k_0).Q'$

 k_0 cannot be free in Q

The Problem with Equating α -equivalent Terms

The idea behind R-PASS:

throw
$$k[k_0]; P \mid \text{catch } k(k_1).Q \rightarrow P \mid Q'$$
 if catch $k(k_1).Q \equiv_{\alpha} \text{catch } k(k_0).Q'$

 k_0 cannot be free in Q

We used a **standard** representation of binders, to mechanise a **non-standard**, **but correct** use of binders.

Mechanising the Revised System

Yoshida, N. and Vasconcelos, V. T. (2007). Language primitives and type discipline for structured communication-based programming revisited: Two systems for higher-order session communication. Electronic Notes in Theoretical Computer Science, 171(4):73 – 93. Proceedings of the First International Workshop on Security and Rewriting Techniques (SecReT 2006)

Processes, Channels, and Polarities

A solution to the naive $_{\text{R-PASS}}$ is to distinguish **channel polarities** ([Gay and Hole, 2005])

$$p,q := + \mid -$$
 throw $k^p \left[k'^q \right]; P \mid \operatorname{catch} k^{\overline{p}} \left(\right). Q \to P \mid Q^{k'^q}$

```
Inductive channel :=
| Ch of (kvar * polarity)
| Var of cvar.

Inductive proc : Set :=
| send : channel → exp → proc → proc
| receive : channel → proc → proc
| throw : channel → channel → proc → proc
| catch : channel → proc → proc
| nu_ch : proc → proc
```

Atoms: Separating Namespaces

```
Module CA := AtomScope Atom.Atom. (* Module of the atoms for channels *)
Module KA := AtomScope Atom.Atom. (* Module of the atoms for channel name *)
Module EA := AtomScope Atom.Atom. (* Module of the atoms for expressions *)

Notation cvar := (CA.var).
Notation kvar := (KA.var).

Notation evar := (EA.var).
```

Typing

```
Inductive tp: Set := 

| input: sort \rightarrow tp \rightarrow tp 

| output: sort \rightarrow tp \rightarrow tp 

| ended: tp | ...
```

```
Inductive oft : sort_env → proc → tp_env → Prop :=
| t_send : forall G kt e P D S T,
    oft_exp G e S →
    oft G P (add kt T D) →
    oft G (send (chan_of_entry kt) e P) (add kt (output S T) D)
| ...
```

Linear Environments

```
\Delta ::= \cdot \mid \Delta, k : \alpha
Inductive env := Undef | Def of {finMap K \rightarrow V}.

Definition add x t E :=

if x \in dom E then Undef else upd x t E.
```

We defined operations on environments that contain **linear** channels.

Adding a channel that is already in the environment results in an **undefined** environment.

We use this fact pervasively in our mechanised proofs.

Subject Reduction

```
Theorem SubjectReduction G P Q D:
oft G P D \rightarrow P \longrightarrow * Q \rightarrow exists D', oft G Q D'.
```

Mechanising a Proof of Subject Reduction

Using separate namespaces requires us to prove distinct **substitution lemmas** for every different kind of binder (expression, channel, shared channel).

Separate namespaces helps us avoid errors (e.g. using a channel instead of an expression), and simplifies proofs.

Linear environments allow us to make simplifying assumptions about defined environments.

Summary of Act I

- Deep Embedding binders allows us to fully control the calculus.
- LN requires a number of theorems and lemmas to prove our basic safety properties.
- EMTST (our tool) helps with nominal sets and environments.
- Next, we will explore what do we gain if we give up control (using shallow embeddings).

- Gay, S. and Hole, M. (2005). Subtyping for Session Types in the Pi Calculus. Acta Informatica, 42(2):191–225.
- Honda, K., Vasconcelos, V. T., and Kubo, M. (1998). Language primitives and type discipline for structured communication-based programming. In Hankin, C., editor, Programming Languages and Systems, pages 122–138, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Yoshida, N. and Vasconcelos, V. T. (2007). Language primitives and type discipline for structured communication-based programming revisited: Two systems for higher-order session communication. Electronic Notes in Theoretical Computer Science, 171(4):73 93. Proceedings of the First International Workshop on Security and Rewriting Techniques (SecReT 2006).

Act II

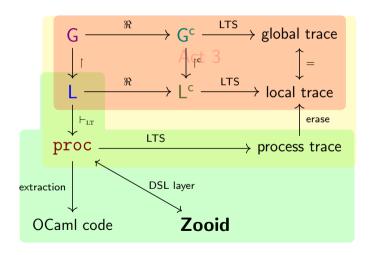
embedding

Smol-Zooid: multiparty with shallower

Goals

- 1. Certifying individual processes of a distributed system
- 2. Extracting runnable code
- 3. Avoiding complex formalisations of binders, whenever possible

Overview



Smol Zooid

- We combine shallow/deep embeddings of binders
 - Processes are defined inductively
 - Values are standard Gallina values
 - We use DeBruijn indices for the deeply embedded binders
- SZooid constructs are well-typed by construction

- We leverage Coq code extraction mechanism
- For simplicity, SZooid does not cover choices

Core Processes

In: http://github.com/emtst/gentleAdventure

Payload Types

We need to define a type for payload types:

- We need a decidable equality on payload types
- We need a decidable equality on payload values

```
Inductive type := Nat | Bool | ...
Definition interp_type : type -> Type := ...
```

Semantics: overview

$$P \xrightarrow{E} P'$$

- What is an event?
- How to manage recursion?

Semantics: events

The semantics is an LTS:

- the labels are the communication events
- it is parameterised by a payload interpretation function
- traces are obtained as the greatest fixpoint of the LTS step

```
Inductive action := a_send | a_recv.
Record event interp_payload :=
    { action_type : action;
        subj : participant;
        party : participant;
        payload_type : type;
        payload : interp_payload payload_type }.
```

Semantics: Recursion Variables

p_unroll exposes the first communication action in a process (unfold recursion):

```
Definition p_unroll e :=
match e with
| Rec e' => p_subst 0 e e'
| e' => e'
end.
```

Semantics: Recursion Variables

```
Fixpoint p_subst d e' e :=
match e with
Rec e => Rec (p_subst d.+1 e' e)
| Jump X => if X == d then p_shift d 0 e' else e
end.
Example ex_p_subst:
p_subst 0 (ping_Alice) (Rec (Jump 1)) = Rec ping_Alice.
```

Semantics: recursion unrolling

```
(* unroll uT. Alice!O. T to Alice!O. uT. Alice!O. T *)
Example ex_p_unroll:
p_unroll (Rec (@Send Alice Nat O (Jump O)))
= @Send Alice Nat O (Rec (@Send Alice Nat O (Jump O))).
```

Semantics: step

The step of the LTS is defined as a **function**:

```
Definition step' e E :=
  match e with
  | Send p T x k =>
    if (action_type E == a_send) && (party E == p) &&
       (eq_payload (payload E) x)
    then Some k else None
  | \text{Recv p T k} => \dots | => \text{None}
  end.
Definition step e := step' (p_unroll e).
```

Semantics: step

```
Definition event_alice: event interp_type :=
{| action_type := a_send;
   from := Bob;
   to := Alice:
   payload_type := Nat;
   payload := 0 \mid \}.
Example ex_step: step infinite_ping_Alice event_alice
                = Some infinite_ping_Alice.
```

Local Types

- Local types for processes
- Notion of "Being well-typed"
- Simultaneous construction of processes & well-typeness proof

Local Types

```
Inductive lty :=
    | l_end
    | l_jump (X : nat)
    | l_rec (k : lty)
    | l_send (p : participant) (T : type) (1 : lty)
    | l_recv (p : participant) (T : type) (1 : lty).
```

Type System

```
Inductive of_lty : proc -> lty -> Prop :=
| lt_Send pTkLx:
   of_lty k L -> of_lty (@Send p T x k) (l_send p T L)
Example ex_of_lty:
of_lty infinite_ping_Alice
(l_rec (l_send Alice Nat (l_jump 0))).
```

Smol Zooid: Smart Constructors

- It would be tedious to type up both a local type and a process
- Users would need to provide a proof that processes are well-typed

We define **SZooid** (Smol Zooid), to write well-typed processes by construction, avoiding repetition.

Smol Zooid: Smart Constructors

Smol Zooid: Smart Constructors

Conclusion

- 1. Define processes and local types
- 2. Semantics of processes
- 3. Automatic construction of local types
 - We also have code extraction
 - and subject reduction

Coq-Metatheory for Smol-Zooid

A Gentle Adventure Mechanising Message Passing Concurrency Systems, Act 3

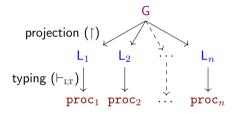
David Castro-Perez, Francisco Ferreira, Lorenzo Gheri, Martin Vassor, and Nobuko Yoshida

DisCoTec 2022 Lucca, 13-06-2022 Imperial College London

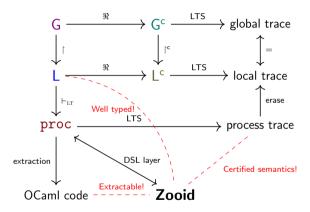




The MPST World, as We Know It

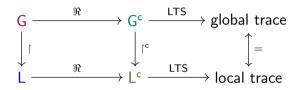


Zooid



D. Castro-Perez, F. Ferreira, L. Gheri, and N. Yoshida. <u>Zooid: a DSL for certified multiparty computation:</u> from mechanised metatheory to certified multiparty processes. <u>PLDI 2021</u>

Introducing the Metatheory of Smol-Zooid Types



- unravelling preserves projection; focus on coinduction (1st square)
- trace equivalence; focus on soundness (2nd square)

https://github.com/emtst/GentleAdventure/act3

Formalisation of Global and Local Types

Inductively Defined Datatypes Coinductively Defined Datatypes

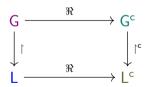
$$\begin{array}{lll} {\sf G} & ::= {\sf end} & & {\sf G}^{\sf c} ::= & {\sf end}^{\sf c} \\ & | \ \mu X.{\sf G} & & | \ p \rightarrow {\sf q} : ({\sf S}).{\sf G}^{\sf c} \\ & | \ p \rightarrow {\sf q} : ({\sf S}).{\sf G}^{\sf c} \\ & | \ p \rightarrow {\sf q} : ({\sf S}).{\sf G}^{\sf c} \\ \end{array}$$

Formalisation of Global and Local Types

Abandoning Inductive Datatypes

Theorem (Unravelling preserves projections)

Given G, L, G^c and L^c, such that (a) G|r = L, (b) G \Re G^c, and (c) L \Re L^c, then G^c |cr L^c.



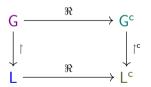
Proof.

By coinduction. :)

Abandoning Inductive Datatypes

Theorem (Unravelling preserves projections)

Given G, L, G^c and L^c, such that (a) G|r = L , (b) G \Re G^c, and (c) L \Re L^c, then G^c |cr L^c.

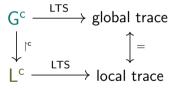


 $\longrightarrow \mathsf{Coq}$

Proof.

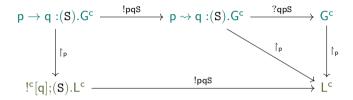
By coinduction. :)

Type Semantics for Zooid

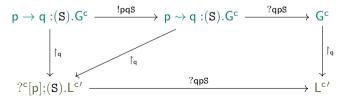


With Love, from p to q

p sends:



q receives:



Tools for our LTS

Actions. !pqS and ?qpS

(Local) Environments. E such that, $E(p) = L^c_p$ where $G^c \upharpoonright^c p L^c_p$

Queues and Queue Environments. Q, buffers for asynchronous communication.

Tools for our LTS

Actions. !pqS and ?qpS

(Local) Environments. E such that, $E(p) = L^c_p$ where $G^c \upharpoonright^c p L^c_p$

Queues and Queue Environments. Q, buffers for asynchronous communication.



Theorems

Theorem (Step Soundness)

If $G^c \xrightarrow{a} G^{c'}$ and $G^c \upharpoonright \upharpoonright (E,Q)$, there exist E' and Q' such that $G^{c'} \upharpoonright \upharpoonright (E',Q')$ and $(E,Q) \xrightarrow{a} (E',Q')$.

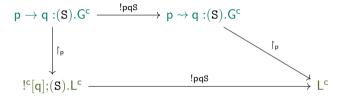
Theorem (Step Completeness)

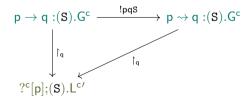
If $(E,Q) \xrightarrow{a} (E',Q')$ and $G^c \upharpoonright (E,Q)$, there exist $G^{c'}$ such that $G^{c'} \upharpoonright (E',Q')$ and $G^c \xrightarrow{a} G^{c'}$.

Theorem (Trace equivalence)

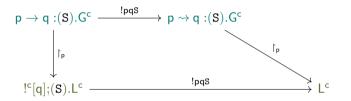
If $\mathsf{G^c} \upharpoonright \upharpoonright (E,Q)$, then $tr^g t\mathsf{G^c}$ if and only if $tr^l t(E,Q)$.

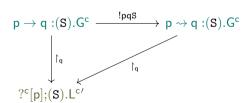
Lemma, to give the flavour





Lemma, to give the flavour







You Suffer...

- Formal proofs are not easy.
- Proof design is the key.
- Proof techniques are to be taken seriously: (co)induction, functions VS relations, treatment of bindings...
- D. Castro-Perez, F. Ferreira, L. Gheri, and N. Yoshida. "Zooid: a DSL for certified multiparty computation: from mechanised metatheory to certified multiparty processes". PLDI 2021.
 DOI: https://doi.org/10.1145/3453483.3454041
 website: http://mrg.doc.ic.ac.uk/publications/zooid-paper/
- → This tutorial is available at https://github.com/emtst/GentleAdventure

... but Why?

Formal proofs are not easy

¹Aydemir et al. "Mechanized Metatheory for the Masses: The POPLmark challenge." 2005

... but Why?

Formal proofs are not easy, but useful and fun!

As witnessed, e.g., by the influential POPLmark Challenge¹...

¹Aydemir et al. "Mechanized Metatheory for the Masses: The POPLmark challenge." 2005

... but Why?

Formal proofs are not easy, but useful and fun!

As witnessed, e.g., by the influential POPLmark Challenge¹...

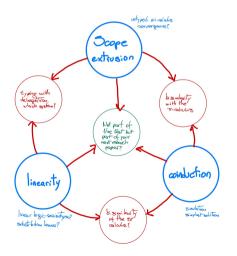
Towards a Concurrent Calculi Formalisation Benchmark

Challenge problems:

- name passing and scope extrusion
- linearity and behavioural type systems
- coinduction and reasoning about process algebras

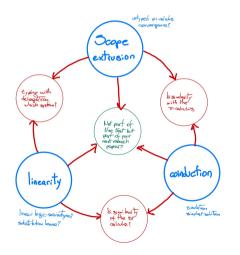
¹Aydemir et al. "Mechanized Metatheory for the Masses: The POPLmark challenge." 2005

The Future is Unwritten... But Sketched!



- Concurrent Benchmark website: https://concurrentbenchmark.github.io/
- This tutorial: https://github.com/emtst/GentleAdventure

The Future is Unwritten... But Sketched!



- Concurrent Benchmark website: https://concurrentbenchmark.github.io/
 - This tutorial: https://github.com/emtst/GentleAdventure

THANK YOU!